There may be more than one way to correctly answer a question, or several ways to describe the same answer to a question. Not all of the possible correct answers are listed here.

1. (a) (4 points) \( \text{Var}(Y) = \sigma^2 I \) and \( \text{E}(Y) = X\beta \)

   (b) (4 points) The models \( Y = X_1\beta_1 + \epsilon \) and \( Y = X_2\beta_2 + \epsilon \) are reparameterizations of each other if there exists a matrix \( G \) such that \( X_2 = X_1G \) and a matrix \( F \) such that \( X_1 = X_2F \). This is equivalent to saying that the space spanned by the columns of \( X_2 \) is the same as the space spanned by the columns of \( X_1 \).

   (c) (4 points) Yes. Use 
   \[
   F = \begin{bmatrix}
   0 & 0 & 0 \\
   1 & 0 & 0 \\
   0 & 1 & 0 \\
   0 & 0 & 1
   \end{bmatrix}
   \quad \text{and} \quad
   G = \begin{bmatrix}
   1 & 1 & 0 \\
   1 & 0 & 1 \\
   1 & 0 & 0
   \end{bmatrix}.
   \]

   (d) (9 points) \( \alpha_1 \) is estimable in model 1 and model 3 because \( X_1 \) and \( X_3 \) have full column rank. Alternatively, you can directly use the definition of estimability. \( \alpha_1 \) is estimable in model 1 because \( \text{E}(Y_1) = \alpha_1 \) and \( \alpha_1 \) is estimable in model 3 because \( \text{E}(Y_1) = \alpha_1 \). For model 2, note that \( X_2a = 0 \), for \( a = (1 \ -1 \ -1 \ -1)^T \) and \( c^T\beta_2 \) is estimable if and only if \( c^T\beta_2 = c_1 \cdot c_2 - c_3 - c_4 = 0 \). For model 2, \( \alpha_1 = (0 \ 1 \ 0 \ 0)\beta_2 \) does not satisfy this condition and \( \alpha_1 \) is not estimable.

   (e) (6 points) \( A = \left(X_2^T X_2\right)^{-1} X_2^T \) for any generalized inverse \( \left(X_2^T X_2\right)^{-1} \). This produces the ordinary least squares estimator, which is also the best linear unbiased estimator of \( \text{C} \beta_2 \), and the value of this estimator \( \text{C}b = C\left(X_2^T X_2\right)^{-1} X_2^T Y \) is unique when \( \text{C} \beta_2 \) is estimable.

   (f) (6 points) The null hypothesis \( H_0: \alpha_1 = \alpha_2 = \alpha_3 \) can be written as 
   \[
   H_0: \text{C} \beta_2 = \begin{bmatrix}
   0 & 1 & -1 & 0 \\
   0 & 0 & 1 & -1
   \end{bmatrix}
   \begin{bmatrix}
   \alpha_0 \\
   \alpha_1 \\
   \alpha_2 \\
   \alpha_3
   \end{bmatrix}
   = \begin{bmatrix}
   0 \\
   0
   \end{bmatrix}.
   \]
   From the answer to part (d), \( \alpha_1 - \alpha_2 \) and \( \alpha_2 - \alpha_3 \) are estimable functions. Furthermore, \( \text{C} \) has row rank equal to 2. Hence, \( H_0: \text{C} \beta_2 = \begin{bmatrix}
   0 \\
   0
   \end{bmatrix} \) is a testable hypothesis.

   (g) (18 points) (i) and (ii) For any solution \( b = \left(X_2^T X_2\right)^{-1} X_2^T Y \) to the normal equations, use 
   \[
   \text{C}b = C\left(X_2^T X_2\right)^{-1} X_2^T Y \text{ to estimate } \text{C} \beta_2 = \begin{bmatrix}
   0 & 1 & -1 & 0 \\
   0 & 0 & 1 & -1
   \end{bmatrix}
   \begin{bmatrix}
   \alpha_0 \\
   \alpha_1 \\
   \alpha_2 \\
   \alpha_3
   \end{bmatrix}
   \]
   and compute the numerator.
sum of squares

\[ SS_{H_0} = b^T C^T \left( C \left( X_2^T X_2 \right) - C^T \right)^{-1} C b = b^T C^T A C b \]

Then, \( \Sigma = \text{Var}(Cb) = \sigma^2 C \left( X_2^T X_2 \right) - C^T \), and we have \( \frac{1}{\sigma^2} \Lambda \Sigma = I \) is an idempotent matrix.

By Result 4.6 in the notes, \( \frac{1}{\sigma^2} SS_{H_0} = \frac{1}{\sigma^2} b^T C^T \left( C \left( X_2^T X_2 \right) - C^T \right)^{-1} C b \) has a chi-square distribution with \( 2 = \text{rank}(C) \) degrees of freedom and non-centrality parameter \( \delta^2 = \frac{1}{\sigma^2} \beta^T C^T \left( C \left( X_2^T X_2 \right) - C^T \right)^{-1} C \beta \). Compute the denominator sum of squares as \( SSE = Y^T (I - P_{X_2}) Y \). Then, \( \Sigma = \text{Var}(Y) = \sigma^2 I \), and we have \( \frac{1}{\sigma^2} (I - P_{X_2}) \Sigma = I - P_{X_2} \) is an idempotent matrix. By Result 4.6 in the notes,

\[ \frac{1}{\sigma^2} SSE = \frac{1}{\sigma^2} Y^T (I - P_{X_2}) Y \] has a central chi-squared distribution with rank(\( I - P_{X_2} \)) = 12 - 3 = 9 degrees of freedom, because the noncentrality parameter is \( \delta^2 = \beta^T X_2^T (I - P_{X_2}) X_2 \beta = 0 \).

(ii) To show that the sums of squares are independent, note the \( SSE \) is a function of \( (I - P_{X_2}) Y \), only, and \( SS_{H_0} \) is a function of \( b = \left( X_2^T X_2 \right)^{-1} X_2^T Y \), only. Since \( Y \sim N(X_2 \beta_2, \sigma^2 I) \), we have \( \left[ \left( X_2^T X_2 \right)^{-1} X_2^T \right] \left[ \left( X_2^T X_2 \right)^{-1} X_2^T \right] Y \) has a multivariate normal distribution with \( \text{Cov} \left( \left( X_2^T X_2 \right)^{-1} X_2^T Y, (I - P_{X_2}) Y \right) = \left( X_2^T X_2 \right)^{-1} X_2^T \left( \sigma^2 I \right) (I - P_{X_2}) = 0 \). Hence, \( SSE \) is distributed independently of \( SS_{H_0} \).

Alternatively, you could have written \( SS_{H_0} \) as \( SS_{H_0} = Y^T (P_{X_2} - P_1) Y \) and made use of Cochran’s theorem to show that \( SSE \) is distributed independently of \( SS_{H_0} \).

(iii) The test is performed by rejecting the null hypothesis if \( F = \frac{SS_{H_0}/2}{SSE/9} > F_{(2,9), \alpha} \). The noncentrality parameter for this test is \( \delta^2 = \frac{1}{2\sigma^2} \beta^T C^T \left( C \left( X_2^T X_2 \right) - C^T \right)^{-1} C \beta \), which is zero if and only if \( H_0: C \beta_2 = \left[ \begin{array}{c} 0 \\ 0 \end{array} \right] \) is true.
2. (9 points) Everyone did well on this question. There are various ways to compute the same thing in S-Plus. One set of code is shown below. This code assumes that you have access to the library containing the ginv() function. It also assumes that the column headings on page 1 of the exam are not included in the data file.

```r
> w <- matrix(scan("herb.dat"),ncol=4,byrow=T)
> x2 <- cbind(1, w[, 2:4])
> px2 <- x2%*%ginv(t(x2)%*%x2)%*$t(x2)
> qr(x2)$rank
> qr(px2)$rank
```

A number of students entered the model matrix directly after entering the herb.dat file, instead of using the second line of the code shown above to create the model matrix. I do not understand why so many students did that.

3. (a) (4 points) Proportional sample sizes \( n_{ij} = \frac{n_{i*}n_{j*}}{n_{**}} \) for all \((i,j)\).

(b) (4 points) \( \alpha_3 = \beta_3 = \gamma_{13} = \gamma_{23} = \gamma_{33} = \gamma_{31} = \gamma_{32} = 0 \)

(c) (4 points) \( \hat{\alpha}_2 \) is an unbiased estimator of \( E(\bar{Y}_{23*}) - E(\bar{Y}_{33*}) = \mu_{23} - \mu_{33} \), the difference between the mean responses for the second and third levels of factor A when factor B is at the third level.

4. (a) (6 points) \( \frac{1}{\sigma^2} R(\gamma \mid \mu, \alpha, \beta) \) has a chi-square distribution with 2 degrees of freedom. The null hypothesis is \( H_0: \gamma_{11} - \gamma_{12} - \gamma_{21} + \gamma_{22} = 0 \) and \( \gamma_{21} - \gamma_{23} - \gamma_{31} + \gamma_{33} = 0 \). Due to the missing cells, some interaction contrasts are not estimable.

(b) (6 points) Use the definition of an estimable function:

\[
\alpha_1 - \alpha_2 = E(\bar{Y}_{11*}) - E(\bar{Y}_{21*}) = \mu_{11} - \mu_{21}
\]

\[
\alpha_1 - \alpha_3 = E(\bar{Y}_{11*}) - E(\bar{Y}_{31*}) = \mu_{11} - \mu_{31}
\]

\[
\alpha_2 - \alpha_3 = E(\bar{Y}_{21*}) - E(\bar{Y}_{31*}) = \mu_{23} - \mu_{33}
\]

(c) (8 points) No. For \( \mathbf{c}_T \mathbf{t} \) to have a chi-square distribution we would need to find a constant \( \mathbf{c}_1 \) such that \( \mathbf{c}_1(\mathbf{A}^T \mathbf{A}) \mathbf{Var}(\mathbf{Y}) = \mathbf{c}_1(\mathbf{A}^T \mathbf{A})(\sigma^2 \mathbf{I}) = \mathbf{c}_1 \sigma^2 (\mathbf{A}^T \mathbf{A}) \) is an idempotent matrix. Note that \( \mathbf{A}^T \mathbf{A} = \\
\begin{bmatrix}
0.25(\mathbf{1}_{2x1}\mathbf{T}\mathbf{1}_{2x1}) & 0_{2x6} & -0.25(\mathbf{1}_{2x1}\mathbf{T}\mathbf{1}_{2x1}) \\
0_{6x2} & \frac{1}{36}(\mathbf{1}_{6x1}\mathbf{T}\mathbf{1}_{6x1}) & -\frac{1}{12}(\mathbf{1}_{6x1}\mathbf{T}\mathbf{1}_{6x1}) \\
-0.25(\mathbf{1}_{2x1}\mathbf{T}\mathbf{1}_{2x1}) & -\frac{1}{12}(\mathbf{1}_{2x1}\mathbf{T}\mathbf{1}_{6x1}) & 0.25(\mathbf{1}_{2x1}\mathbf{T}\mathbf{1}_{2x1}) \\
\end{bmatrix}
\]
$A^T AA^T A = \begin{bmatrix}
.25 (l_{2 \times 1} l_{2 \times 1}) & \frac{1}{24} (l_{2 \times 1} l_{6 \times 1}) & -\frac{1}{18} (l_{6 \times 1} l_{2 \times 1}) \\
-\frac{1}{24} (l_{2 \times 1} l_{6 \times 1}) & \frac{1}{54} (l_{6 \times 1} l_{2 \times 1}) & -\frac{1}{18} (l_{6 \times 1} l_{2 \times 1}) \\
-\frac{1}{25} (l_{2 \times 1} l_{2 \times 1}) & -\frac{1}{18} (l_{2 \times 1} l_{6 \times 1}) & \frac{7}{24} (l_{6 \times 1} l_{2 \times 1}) 
\end{bmatrix}$

Hence, $A^T AA^T A$ is not a multiple of $A^T A$, and there is no way to select $c_1$ to make $c_1 \sigma^2 (A^T A)$ idempotent. Consequently, no multiple of $T$ has a chi-square distribution.

(d) (8 points) Reject the null hypothesis if $F = \frac{R(\alpha, \beta) / 2}{Y^T (I - P_X) Y / 5} > F_{(2, 5), \alpha}$. Equivalently, you could perform the same F-test by rejecting the null hypothesis if

$$F = \frac{b^T C^T \left[ C (X^T X)^{-1} C^T \right]^{-1} C b}{Y^T (I - P_X) Y / 5} > F_{(2, 5), \alpha} \quad \text{where} \quad C \beta = \begin{bmatrix}
\alpha_0 \\
\alpha_1 \\
\alpha_2 \\
\alpha_3 \\
\beta_1 \\
\beta_2 \\
\beta_3
\end{bmatrix}.$$

The exam scores (out of 100 points) are presented as a stem-leaf display:

<table>
<thead>
<tr>
<th>Score</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>90</td>
<td>67</td>
</tr>
<tr>
<td>90</td>
<td>00003</td>
</tr>
<tr>
<td>80</td>
<td>7999</td>
</tr>
<tr>
<td>80</td>
<td>00122233</td>
</tr>
<tr>
<td>70</td>
<td>7799</td>
</tr>
<tr>
<td>70</td>
<td>00034</td>
</tr>
<tr>
<td>60</td>
<td>567799</td>
</tr>
<tr>
<td>60</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>68</td>
</tr>
<tr>
<td>50</td>
<td>3</td>
</tr>
<tr>
<td>40</td>
<td>8</td>
</tr>
<tr>
<td>40</td>
<td>3</td>
</tr>
</tbody>
</table>